



---

## LA COLABORACIÓN CIENTÍFICA / SCIENTIFIC COLLABORATION

---

### Acceso abierto a los datos de investigación, una vía hacia la colaboración científica

Remedios Melero\*, Javier Hernández-San-Miguel\*\*

\*Instituto de Agroquímica y Tecnología de Alimentos- CSIC. Correo-e: [rmelero@iata.csic.es](mailto:rmelero@iata.csic.es)

\*\*Universitat Politècnica de València. Biblioteca. Correo-e: [frahersa@bib.upv.es](mailto:frahersa@bib.upv.es)

Recibido: 03-02 -2014; 2ª versión: 14-04-2014; Aceptado: 25-04-2014.

**Cómo citar este artículo/Citation:** Melero, R.; Hernández-San-Miguel, J. (2014). Acceso abierto a los datos de investigación, una vía hacia la colaboración científica. *Revista Española de Documentación Científica*, 37(4): e066. doi: <http://dx.doi.org/10.3989/redc.2014.4.1154>

**Resumen:** La transmisión del conocimiento debería verse favorecida por las oportunidades que brindan las TIC como medio de acceso y distribución de objetos digitales. Sin embargo, las barreras que impiden el acceso y reutilización de los trabajos derivados de la actividad científica y académica, ya sean económicas o de derechos de explotación, inhiben compartir un bien común como es el conocimiento. El movimiento por el acceso abierto a la ciencia promueve la eliminación de estas barreras y aboga por una cultura que permita compartir y reutilizar materiales, siempre con el reconocimiento de la autoría y con un uso responsable. Si el artículo científico ha sido históricamente una forma esencial de la comunicación de la ciencia, en la era digital cobran relevancia sus fundamentos, entre ellos los datos observacionales, descriptivos o experimentales que subyacen al artículo. Los datos pueden reutilizarse, transformarse mediante nuevos métodos o agregarse a otras fuentes. Los datos en abierto evitan la duplicidad de ensayos, dan transparencia a su forma de obtención y permiten su validación. En este trabajo se presentan algunas iniciativas y recomendaciones de cómo compartir, gestionar y promocionar el acceso abierto a los datos generados durante la investigación científica, como vía de colaboración entre grupos o personas con afinidad en sus temas de trabajo.

**Palabras clave:** Acceso abierto; datos; repositorios de datos; citación de datos; políticas de acceso abierto.

#### Open access to research data: a track towards scientific collaboration

**Abstract:** Knowledge transfer should be facilitated by the opportunities offered by information technologies, as they affect the access and distribution of digital objects. However barriers to the access and reuse of scholarly research, whether economic or copyright-related, inhibit the sharing of this valuable common good. The open access movement promotes the elimination of these barriers and advocates for an open access culture of sharing and reusing materials, while guaranteeing that authors be duly acknowledged and that the data be used responsibly. If scientific papers have historically been essential for the communication of science, in the digital age it is now their building blocks that have gained greater importance, especially the observational, descriptive or experimental data that underpin the articles. Open research data can be reused, transformed by new methods or aggregated to other sources. Open access to research data avoids redundancy, provides transparency on how they have been obtained and allow for their validation. This work provides an overview of some initiatives and recommendations on how to share and manage research data and foster open access to it, as a means of collaboration between groups or individuals working in similar disciplines.

**Keywords:** Open access; research data; data repositories; data citation; open access policies.

**Copyright:** © 2014 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

## 1. INTRODUCCIÓN

El flujo de la comunicación científica desde la aparición de las primeras revistas científicas en el siglo XVII ha seguido más o menos el mismo orden: creación, evaluación, certificación, distribución y preservación. Esta estructura implica a creadores y evaluadores, a las editoriales que certifican y distribuyen los trabajos y a las bibliotecas que se ocupan de la mayor parte de la preservación. La revolución digital, Internet y la multiplicidad de aplicaciones desarrolladas en los últimos años han hecho que esta cadena amplíe sus formas de distribución y difusión del conocimiento, a lo que ha contribuido el movimiento por el acceso abierto a la ciencia, más conocido como movimiento *Open Access*.

Las primeras declaraciones, Declaración de Budapest, de Bethesda y de Berlín, sientan las bases del acceso abierto a la ciencia, y al hablar de contribuciones al conocimiento incluyen materiales muy diversos como son los trabajos de la investigación científica, datos primarios y metadatos, materiales gráficos y pictóricos, materiales académicos multimedia y material docente, entre otros. Sin embargo, parece que el acceso abierto en algunos sectores se haya centrado principalmente en conseguir que los trabajos publicados en revistas científicas estén en acceso abierto. No obstante, la tendencia a tratar a las fuentes, sean de la naturaleza que sean, como objetos digitales enriquece la diversidad de ese universo disponible a través de Internet, que incluye los trabajos derivados de la investigación y de la actividad académica de profesores e investigadores: datos, material docente, informes, libros, entre otros.

De acuerdo con las declaraciones mencionadas en favor del acceso abierto, las barreras que inhiben el acceso abierto son las económicas y las derivadas de los derechos de explotación de las obras: reproducción, distribución, comunicación pública y transformación. El autor o creador de la obra es el poseedor de estos derechos, salvo que los ceda a terceros, caso muy frecuente cuando se trata de trabajos publicados en revistas, en las que, mediante la cesión exclusiva, los autores pierden el control de los mismos, generalmente por la firma del *Copyright Transfer Agreement*. Por eso, el uso de licencias de publicación o de distribución, que permitan una reutilización de los trabajos con fines responsables, facilita el acceso abierto a los correspondientes objetos digitales, tanto a través de repositorios institucionales o temáticos de acceso abierto como a través de las revistas que cumplan con las condiciones de acceso abierto, en mayor o menor grado (SPARC, 2014).

El acceso abierto implica acceso gratuito y permisos para la reutilización de las obras, acceder gratis a las fuentes originales no es suficiente, aunque "es mejor que nada". El acceso abierto a la ciencia no sólo significa acceder sin barreras a la literatura científica, significa poder compartir y reutilizar

esos recursos generados mayoritariamente de proyectos financiados con fondos públicos. El acceso restringido detiene o desacelera el ciclo de generación de conocimiento.

El acceso abierto significa transparencia, democracia y visibilidad, y también rentabilidad porque permite un uso eficiente de los recursos, además, el compartir reduce esfuerzos, evita duplicidades, ahorra tiempo e induce a la colaboración. En resumen, el acceso abierto es algo más que acceso, es una nueva visión de concebir la generación y transferencia del conocimiento, no sólo dirigido a la comprensión por humanos, sino para permitir un diálogo entre máquinas mediante sistemas interoperables. En este contexto el concepto de abierto se expande, el instrumento de comunicación no se restringe a los "papers", los datos importan y adquieren relevancia en el contexto de la e-ciencia.

Según la definición de conocimiento abierto y la guía de licencia para datos abiertos de la *Open Knowledge Foundation* (2014a; 2014b), los datos obtenidos de una investigación en abierto (*Open Research Data*) son aquellos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, siempre que se reconozca la autoría y en caso de transformación o reutilización a compartir los nuevos de la misma manera. El acceso a estos datos evita duplicar esfuerzos en su obtención, además su agregación a otros resultados o la aplicación de nuevos métodos para su transformación pueden generar nuevos datos y favorecer otras interpretaciones y colaboraciones interdisciplinarias, tal y como señaló la vicepresidenta de la Comisión Europea, Nelly Kroes (2013), en un discurso sobre *Big Data*:

*"El libre acceso a los resultados científicos y a los datos experimentales es una gran manera de impulsar la ciencia, de impulsar la economía y de permitir que surjan nuevas técnicas y colaboraciones entre disciplinas. Realmente es bastante simple: se trata de asegurar que puedan verse los resultados generados de proyectos financiados mediante el pago de los impuestos..."*

El acceso, la tecnología, la interoperabilidad, la sostenibilidad y la preservación son factores imprescindibles para construir un entorno de e-ciencia colaborativa, accesible hoy y en el futuro. A estos factores hay que sumarles las políticas (institucionales, gubernamentales, de asociaciones profesionales, de empresas...) que promuevan e incentiven la participación en la creación de este modo de concebir la ciencia. No obstante, las políticas, las leyes, las recomendaciones y las directrices ayudan, pero no bastan, crear una cultura de cambio encaminada al acceso abierto en todas sus facetas requiere la participación y colaboración de todas las partes implicadas.

En este trabajo se presentan algunas iniciativas y recomendaciones de cómo compartir, gestionar y promocionar el acceso abierto a los datos genera-

dos durante la investigación científica, como vía de colaboración entre grupos o personas con afinidad en sus temas de trabajo.

## 2. DATOS DE INVESTIGACIÓN EN ACCESO ABIERTO, UNA VÍA DE COLABORACIÓN

Los investigadores, que buscan la mayor visibilidad e impacto de sus publicaciones (Harnad y Brody, 2004), han hecho de las redes sociales de investigación una vía imprescindible para la disseminación de sus investigaciones. Es aquí donde el acceso abierto deriva de un concepto de ciencia abierta hacia uno más complejo de ciencia colaborativa, desde el cual se comparte la investigación y los resultados de la misma (Merlo Vega y otros, 2011). La utilización de plataformas como Academia.edu o ResearchGate, donde los investigadores comparten sus publicaciones y establecen vínculos y relaciones entre ellos son habituales en las estrategias de difusión de su investigación.

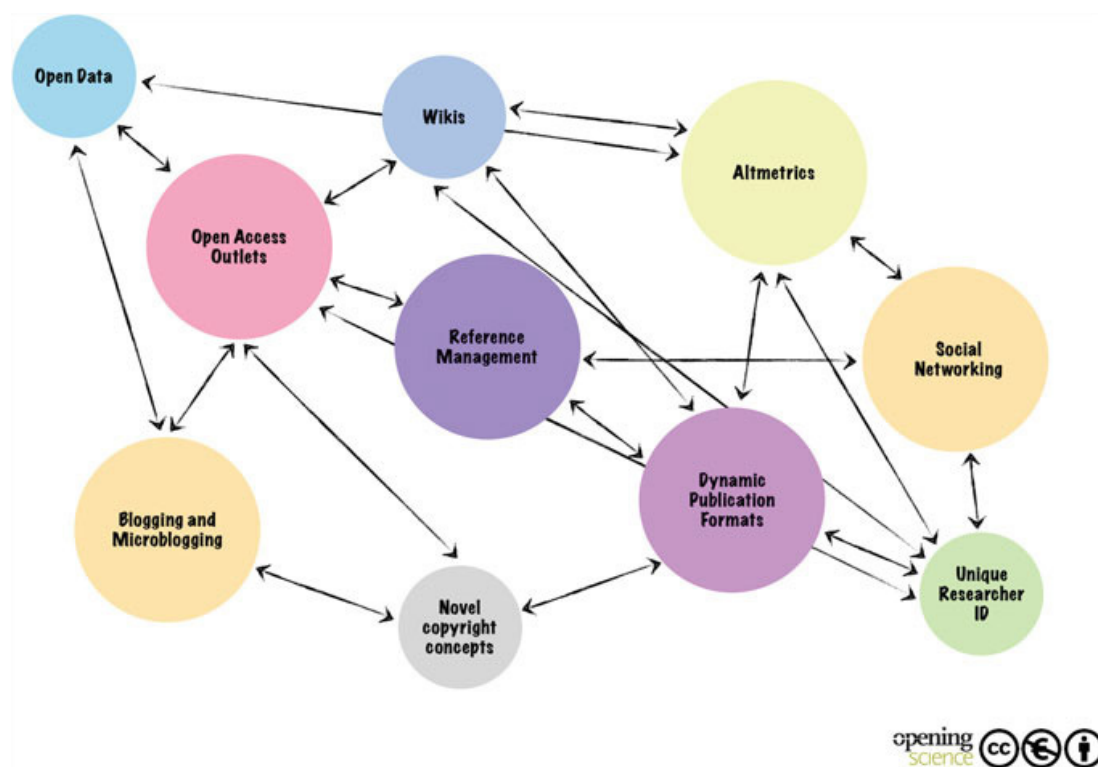
Ciencia 2.0 y Ciencia abierta, aunque no siempre confluyen, se articulan alrededor de las nuevas formas de difusión del conocimiento, donde el acceso abierto a la producción científica se integra con los sistemas tradicionales de la transmisión de la información científica y se desarrollan nuevos flujos para los trabajos de investigación (Figura 1).

La difusión de la producción científica es la fase final del proceso de la investigación científica. En ese sentido la Agenda Digital para Europa indica que "el acceso abierto proporciona a investigadores, empresas y ciudadanos libre acceso en línea a los resultados de investigación financiados por la UE, incluyendo datos de investigación y publicaciones científicas". El objetivo es estimular la cultura de intercambio de datos entre los investigadores y facilitar la reutilización de la información y la ciencia basada en datos de investigación (COAR, 2013).

El tratamiento de los datos de investigación es complejo y requiere de una variedad de soluciones para su tratamiento, gestión y acceso, que permitan su validación y reutilización. Esto implica diferentes estrategias para la integración de la gestión de los datos de investigación en los flujos de trabajo de los investigadores (Ribeiro y Hernández-San-Miguel, 2013). Para ello es necesario la implementación de un plan de gestión de los datos adaptado a la tipología de los datos y que incluya las distintas políticas de acceso, archivo, distribución, reutilización y preservación de los mismos (National Science Foundation, 2013).

Los investigadores toman conciencia del valor de los datos y de su potencial para el reconocimiento de su trabajo, tanto en sus círculos profesiona-

**Figura 1.** Ciencia 2.0 y las nuevas formas de difusión del conocimiento (fuente original: Bartling S. y Friesike S., 2014)



les, como por el aumento en la citación que puede generar su consulta y reutilización. A su vez, mediante el depósito de los datos procedentes de la investigación en un repositorio digital los hacemos accesibles y reutilizables por otros investigadores. Este depósito se puede realizar en repositorios institucionales de datos, en los servicios de datos proporcionados por las entidades financiadoras de la investigación y, en un modelo que se ha vuelto mayoritario, en repositorios de datos temáticos.

### Tipología de datos de investigación. Depósito y localización

Los datos de investigación podemos clasificarlos en función de su obtención o de la metodología aplicada para ella (Borgman, 2012a y 2012b; FECyT, 2012; MIT Libraries, 2014):

- Datos observacionales: registros históricos, se pueden obtener únicamente en un lugar y en un momento en el tiempo (son irrepetibles).
- Datos experimentales: datos que acompañan a los experimentos desde su planificación y preparación hasta la obtención de resultados. Con frecuencia reproducibles pero con un alto coste.
- Datos computacionales, compilados y derivados: datos que suelen incluir datos de entrada, ciertos programas y resultados. Reproducibles con soporte tecnológico.
- Datos de simulación: datos generados a partir de modelos de prueba donde con frecuencia el modelo de entrada es más representativo que la propia salida de datos.

Para poder acceder a los *datasets* habitualmente se depositan en repositorios de datos diseñados con esta finalidad. Tanto las instituciones, los organismos públicos de financiación, las universidades, las fundaciones, las publicaciones periódicas, etc., deberían disponer de los servicios para dirigir a los investigadores y autores en cómo depositar y, a su vez, preservar sus datos de investigación.

El panorama internacional presenta diferentes patrones de repositorios de datos de diferentes tipologías (Coates y otros, 2013): editor (Dryad), consorcio (3TU.Datacentrum), instrumentales (CHANDRA), disciplinares (RKMP), nacionales (Research Data Australia), institucionales (PURR), de propósito general (FigShare), etc.

Podemos plantearnos utilizar cualquiera de los repositorios ya existentes, principalmente disciplinares, o adaptar nuestro repositorio institucional e incluso implementar un repositorio específico para datos. Al proveer al repositorio institucional de una estructura que permita el archivo de los datos conduce a vincular los documentos digitales presentes en los repositorios con sus *datasets*. Algunos ejemplos significativos de esta estrategia institucional los encontramos tanto a nivel internacional (p.e., *Edinburgh Research Archive*, *Universi-*

*ty of Southampton: ePrints Soton*) como nacional (p.e., Digital CSIC). La otra opción sería el diseño de un nuevo repositorio de datos institucional independiente de cualquier otro como en el caso del *Purdue University Research Repository* (PURR) o, especialmente, en el de la *Universidade do Porto* donde van a coexistir tres repositorios: temático (principalmente patrimonial), abierto (con la estructura tradicional de los repositorios institucionales) y de datos (pendiente de implementación).

Entre los repositorios de datos temáticos destacan aquellos dedicados a estructuras proteínicas, secuencias genéticas, de cristalografía y relacionados con temas de ciencias de la tierra, todos ellos de disciplinas en las que los miembros de la comunidad científica a la que pertenecen tienen una alta tradición en compartir datos. De acuerdo con el directorio OpenDOAR existen 27 repositorios temáticos dedicados exclusivamente a *datasets*, el resto hasta 109 son repositorios institucionales (datos a fecha de abril de 2014), gubernamentales o agregadores.

Complementariamente a los registros tradicionales de repositorios (ROAR, Registry of Open Access Repositories y OpenDoar, Directory of Open Access Repositories) alrededor de los datos han surgido iniciativas para intentar normalizar y dar visibilidad a los distintos repositorios de datos que se han creado en los últimos años. De entre los existentes podemos destacar varios especialmente significativos:

*Odisea*: Proyecto español para el inventario internacional de los depósitos que admiten conjuntos de datos de investigación a escala mundial. Permite buscar y sugerir nuevos bancos de datos. Organizado por tipos de materiales, formato y disciplina (Odisea, 2014).

*Databib*: Iniciativa americana con base en Purdue University y vínculos con DataCite (2014). Es una herramienta para ayudar a identificar y localizar los repositorios en línea de datos de investigación. Cuenta con *Advisory Board* y *Editorial Board* de carácter internacional. Es un catálogo de búsqueda, registro, directorio, bibliografía de repositorios de datos de investigación. Los registros describen repositorios de datos que son creados por una comunidad de usuarios, productores de datos, investigadores y bibliotecarios. Es posible el volcado de registros de forma dinámica en formato RDF/XML. Cada registro en Databib expone *linked data* en forma de RDF que está incrustado dentro de la página web que representa a cada repositorio. Todo el registro de metadatos se expresa mediante Dublin Core, FOAF, Creative Commons y utiliza el vocabulario de términos de la Library of Congress (Subject Headings, LSCH).

*Re3data*: Iniciativa alemana financiada por German Research Foundation (DFG). Permi-

te búsquedas por materia, contenido y país. Ofrece información sobre licencias, políticas y ofrece información sobre distintos identificadores persistentes (DOI, URN, ARK, Handle, Purl). Utiliza un schema XML propio (Pampel y otros, 2013).

Recientemente se ha anunciado la futura fusión (2015) de Databib y Re3data, en un único servicio consolidado para evitar la duplicación y confusión de la existencia de dos registros paralelos. El proyecto de unificación está auspiciado por DataCite (2014), el cual asigna Digital Object Identifiers (DOI) a los datasets que almacena.

### Los datos en el ciclo de la citación de trabajos

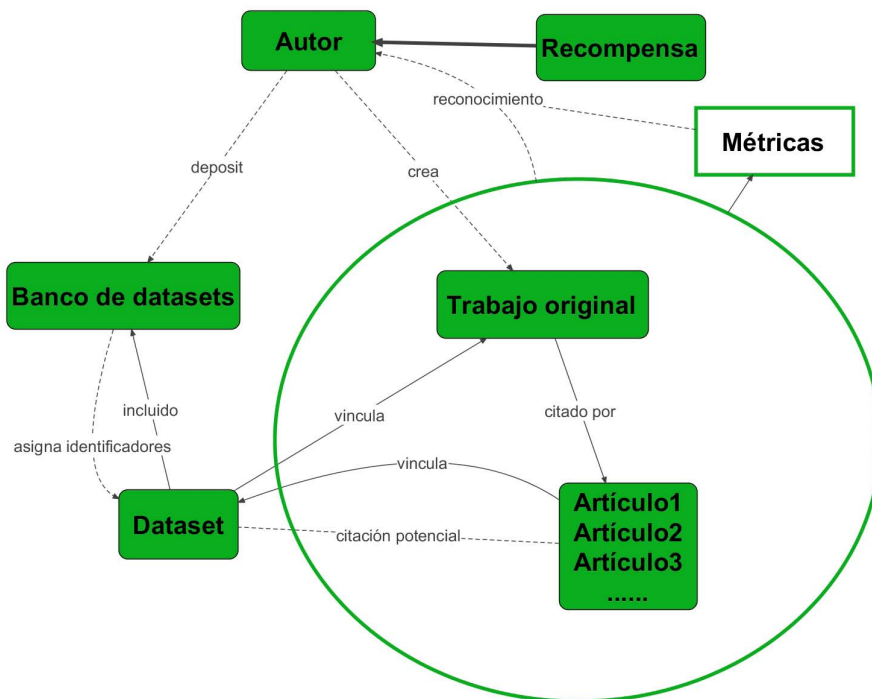
El aumento de la visibilidad de los artículos en abierto potencia o favorece su citación, lo que en términos anglosajones se conoce como *open access citation advantage* (Wagner, 2010; Swan, 2010; Hitchcock, 2013) esto mismo puede ocurrir con los datos. Los ficheros de datos pueden ser citados al igual que cualquier otro objeto digital, pueden ser parte de una publicación o tener identidad propia y visibilidad a través de un repositorio de datos. En ambos casos pueden disponer de su propia licencia de uso y de un identificador único, lo que facilita su potencial citación. Estos *datasets* pueden entrar en el círculo de las citas de forma individual o a través del artículo al que están vinculados (Figura 2). Si bien el tema de las citas está resuelto con los trabajos, todavía no hay consenso sobre la métrica de los *datasets*.

El *Digital Object Identifier* (DOI) también se utiliza como identificador para los ficheros de datos, por ejemplo, como ya se ha mencionado, DataCite asigna DOIs a los *datasets* y expone sus metadatos a través de un proveedor OAI-PMH. Otro proyecto todavía en desarrollo, ODIN, tiene por objetivo vincular los *datasets* a los que ya se les ha asignado un DOI con los autores que los han obtenido a través del identificador ORCID, así se facilita la desambiguación para los autores y se trazan los vínculos con sus datos.

Existen trabajos en los que se ha puesto de manifiesto cómo el acceso a los datos de la investigación, bien a través de la propia revista o bien por la referencia al fichero depositado en un repositorio de datos hace que aumente su citación. Por ejemplo, la revista *Paleoceanography* recibió una media de un 35% más de citas (según datos ISI-Thomson Reuters) en aquellos artículos con enlaces a sus *datasets* depositados en PANGEA. Esa diferencia todavía fue mayor (50%) cuando el análisis se hizo con la revista *The Astrophysical Journal* entre 2000-2010 (Dorch, 2012) con artículos con datos vinculados al ADS (Astrophysical Data System). En otro estudio con trabajos que habían generado secuencias genéticas entre 2000 y 2009 se obtuvo que el 9% de artículos con datos vinculados recibían mayor número de citas (Piwowar y Vision, 2013).

Existen modelos innovadores de revistas, que no solo citan o enlazan sus textos a los *datasets* correspondientes, sino que los datos son su eje cen-

Figura 2. Esquema del ciclo de citación, incluyendo los datos de investigación



tral, como la *Journal of Open Public Health Data* que publica artículos que describen *datasets* que se encuentran en abierto con alto potencial de ser reutilizados. Otra revista de acceso abierto, la *F1000Research*, solicita el número de registro de los datos en un repositorio (*accession number*) para incluirlo en sus textos, y ofrecen a sus autores ayuda para encontrar el repositorio de datos adecuado para depositar sus ficheros. *GigaScience*, nacida en 2012, es otro ejemplo de una revista de acceso abierto que publica estudios sobre *Big Data* obtenidos de experimentos en ciencias de la vida, además ofrece servicios para el archivo de grandes bases de datos en la nube. El grupo editorial *Nature* anunció una nueva publicación similar, llamada *Scientific Data*, para la primavera del 2014. Su objetivo es la publicación de un nuevo tipo de artículo denominado *Data Descriptor*, en el que se describen grandes bases de datos y los procedimientos de su obtención (*Scientific Data*, 2014).

Facilitar los datos también puede conducir a la detección de errores no necesariamente intencionados, sin embargo existe cierta reticencia por parte de los autores a compartir los datos, lo que induce a pensar que quizás haya otros motivos. Por ejemplo, en el análisis de la detección de errores estadísticos de trabajos publicados en unas revistas de psicología se observó que los casos con mayor número de errores fueron aquellos en que los autores no quisieron facilitar los datos de los ensayos (Wicherts y otros, 2011).

### Hacia una cultura de compartir datos

En la figura 3 se muestran de forma esquemática algunos de los beneficios que se obtienen de compartir, reutilizar o transformar datos, basados en el manifiesto *Denton Declaration: An Open Access Data Manifesto* (2012) que se emitió durante la celebración del tercer simposio sobre acceso abierto de la *University of North Texas*. En este manifiesto también se señala la importancia de la preservación, de la sostenibilidad y de las infraestructuras necesarias para la gestión y mantenimiento de los datos de la investigación. Otro factor importante en la creación de una cultura de compartir datos son los roles que juegan todas las partes implicadas en la generación, difusión y promoción de los datos en abierto.

Otros aspectos muy importantes y señalados en diversas fuentes (The Royal Society, 2012) consisten en cómo reconocer, incentivar y agradecer que el autor disponga los datos libremente para usar y compartir, y que todas las partes implicadas, los científicos, la universidad, las asociaciones profesionales, editoriales y el sector privado, de alguna manera contribuyan al cambio por una cultura de datos abiertos (Figura 4).

En un trabajo publicado recientemente (Costas y otros, 2013) se indican unas recomendaciones dirigidas a las diferentes partes que intervienen en la promoción de las posibles ventajas de compartir e intercambiar datos. Que de forma resumida se centran en:

**Figura 3.** Esquema basado en los principios de la *Denton Declaration: An open access manifesto* (<http://openaccess.unt.edu/denton-declaration>)



**Figura 4.** Partes implicadas y acciones que promuevan la cultura de los datos de investigación



*Agencias que financian la investigación*

- Recompensar las actividades de intercambio de datos
- Considerar las métricas de datos en las evaluaciones
- Dar a conocer la importancia y los beneficios del intercambio de datos
- Promover el acceso libre de los datos

*Instituciones de Investigación*

- Promover políticas de intercambio de datos
- Promover los argumentos y los incentivos para compartir datos
- Proporcionar opciones y alternativas para los diferentes tipos de actividades de intercambio de datos
- Profesionalizar el personal y estandarizar las actividades de intercambio de datos (recogida, curación, difusión)

*Científicos*

- Incluir el intercambio de datos como una buena práctica científica y académica
- Promover la citación de datos como la manera formal del reconocimiento por compartir datos

- Llevar a cabo más investigaciones sobre los beneficios y las posibilidades de intercambio de datos
- Definir los códigos de conducta teniendo en cuenta la privacidad y el anonimato, cuando corresponda

*Bibliotecas*

- Promover las publicaciones de datos y las citas de datos
- Formación a los profesores e investigadores para la publicación y citación de datos
- Informar a los autores sobre otros grupos de interés en común de datos (por ejemplo, los proveedores de fondos, depósitos, centros de datos)
- Desarrollar herramientas para encontrar repositorios de datos
- Desarrollar y probar las métricas

*Los centros de datos*

- Informar a la comunidad científica acerca de sus actividades y servicios
- Contribuir a reducir la dispersión de los repositorios de datos
- Desarrollar soluciones sólidas para la preservación y la normalización del almacenamiento de datos y citas
- Desarrollar herramientas para el seguimiento de los usuarios de los repositorios

### Editores

- Promover el intercambio de datos en sus publicaciones y revistas
- Informar a los autores sobre otros agentes interesados en el intercambio de datos (por ejemplo, centros de repositorios de datos)
- Apoyar el acceso abierto a los datos

### Políticas de acceso abierto a los datos de investigación

Existen políticas que requieren o recomiendan el acceso abierto a la ciencia, emitidas tanto de instituciones universitarias y de investigación como de entidades que financian proyectos de investigación, en cuyas convocatorias o bases se especifica de que manera debe favorecerse el acceso abierto a la producción científica. Estas políticas recaen generalmente sobre los artículos publicados, comunicacio-

nes a congresos, informes y tesis doctorales, entre otros. Sin embargo, el requisito o recomendación para hacer que los datos de la investigación queden disponibles en abierto para su consulta y potencial uso no es tan frecuente, tan solo un 26% de las agencias que financian la investigación requieren el depósito de los *datasets* en repositorios de acceso abierto, un 10% lo recomiendan y el resto no hacen alusión al depósito de los datos (datos de SHERPA/JULIET a 29 de enero de 2014). En la tabla 1 se enumeran algunas de estas instituciones que requieren que los datos generados de la investigación sean depositados en los repositorios de acceso abierto para compartir y reutilizar (datos del directorio MELIBEA 2014). Existen más políticas de acceso abierto de universidades y centros de investigación en comparación con las entidades financiadoras; sin embargo respecto a los datos, son estas últimas las que más demandan planes para el archivo y gestión de los datos derivados de los proyectos de investigación.

**Tabla I.** Instituciones con políticas de acceso abierto que incluyen los datos de investigación

<b>Instituciones que financian proyectos de investigación</b>
Canadian Institutes of Health Research
Cancer Research UK
Chief Scientist Office, Scottish Executive
Comunidad Autónoma de Madrid
Consultative Group on International Agricultural Research
Economic and Social Research Council (UK)
Engineering and Physical Sciences Research Council
European Commission. Horizon 2020
European Research Council
Fondazione Cariplo
Fonds zur Förderung der Wissenschaftlichen Forschung
Genome Canada
Gordon and Betty Moore Foundation
Heart and Stroke Foundation
Howard Hughes Medical Institute
Irish Higher Education Authority
Medical Research Council
Michael Smith Foundation for Health Research
Ministero dell'Istruzione, dell'Università e della Ricerca
National Institutes of Health
Natural Environment Research Council
Nederlandse Organisatie voor Wetenschappelijk Onderzoek
Ontario Institute for Cancer Research
Research Councils UK
Science Foundation Ireland
The World Bank
UK Government. Department for International Development
Vetenskapsrådet [Swedish Research Council]
Wellcome Trust
White House Directive
<b>Universidad o centro de investigación</b>
Fraunhofer-Gesellschaft
Koninklijke Nederlandse Akademie van Wetenschappen
National Center for Atmospheric Research
Universidad de Extremadura
Universidad Rey Juan Carlos
Università degli Studi di Torino
Université de Liège



Tampoco es casualidad que sean las entidades que financian proyectos en ciencias de la salud, las que tengan más políticas sobre datos, son sobre todo porque compartir datos de esta naturaleza, siempre respetando la privacidad y confidencialidad de los datos donde corresponda, permite un intercambio de resultados a veces difíciles o imposibles de obtener. En este sentido y auspiciado por la Wellcome Trust, se reunieron representantes de entidades de todo el mundo que financian proyectos relacionados con la salud y emitieron un manifiesto en el que se comprometían a llevar a cabo acciones que fomentaran y garantizaran el compartir los datos de la investigación para el avance de la ciencia, para un mejor aprovechamiento de los recursos y en beneficio de la calidad de la ciencia (Wellcome Trust, 2011).

De las políticas señaladas en la tabla I, cabe destacar la directiva del presidente de los Estados Unidos por la que se requiere establecer un plan para que las publicaciones y los datos generados de la investigación financiada con fondos públicos estén públicamente disponibles, con un embargo máximo de 12 meses (White House Directive, 2013). De esta manera al menos 19 agencias federales están sujetas a la directiva, entre ellas el National Institute of Health, aunque esta última ya tenía su propia política (NIH, 2008). Con respecto a Europa, en el Séptimo Programa Marco se hizo un experimento piloto con proyectos financiados por la CE en las áreas de energía, medio ambiente, salud, tecnologías de la información, la ciencia en la sociedad, infraestructuras, para la investigación, socio-economía y humanidades por el cual las publicaciones derivadas de estos proyectos debían depositarse en un repositorio y estar disponibles en acceso abierto (European Commission, 2009). En el actual programa marco (2014-2020), Horizon 2020, este requisito se ha extendido a todas las disciplinas y ahora el ensayo piloto se aplica a los datos de investigación en las siguientes áreas: tecnologías emergentes, infraestructura para la investigación, tecnologías de la información y comunicación, energías limpias y seguras, cambio climático, medio ambiente, Europa en un mundo cambiante, ciencia para y por la sociedad (European Commission, 2013). Tanto para el Séptimo programa Marco como para el Horizon 2020 ambas iniciativas piloto se apoyan en los proyectos europeos OpenAIRE y su descendiente OpenAIREPlus.

### 3. CONSIDERACIONES FINALES

En este apartado nos gustaría hacer mención e hincapié de dos documentos que extractan las claves de la evolución en la concepción del acceso abierto a la ciencia y que acciones pueden llevarse a cabo para que el *Open Access* sea una realidad y una práctica en nuestra vida académica y profesional.

El primero de ellos es el emitido con motivo del décimo aniversario de la Declaración de Budapest. Se emitió un nuevo manifiesto en el que se abogaba por el "acceso abierto por defecto" (Budapest Open Access Initiative, 2012). El documento creado

y avalado por los padres de la declaración original, ratifica el paradigma del acceso abierto y propone un nuevo reto, es que el acceso abierto a la producción científica se convierta en acceso por defecto en los próximos 10 años. En la Declaración se hacen una serie de recomendaciones para que las instituciones académicas, las entidades financiadoras, los autores y los gestores de la información lo hagan posible. El establecimiento de políticas institucionales y gubernamentales en favor del acceso abierto a la ciencia, la flexibilidad de las editoriales para la gestión de los derechos de explotación, el uso de licencias de distribución abiertas, la promoción y diseminación de la filosofía de lo abierto por defecto constituyen los ejes centrales de esta Declaración.

El segundo documento que nos gustaría reseñar y enfatizar sus conclusiones es un informe elaborado por los socios del proyecto europeo e-InfraNet que lleva por título '*Open' as the Default Modus Operandi for Research and Higher Education* (e-InfraNet, 2013). En él se establecen varias categorías de "abierto", no solo de contenidos sino de infraestructura y de procesos (figura 5).

La cohesión de estos tres pilares: contenido, procesos e infraestructura de componentes "abierto" ("opens") conduce a un cambio cultural. Pero si se produce la desfragmentación en los "opens" puede conducir a la toma de decisiones buenas para unos, que podrían ser adversas para otros, además la sostenibilidad sería más arriesgada y el aprovechamiento de los recursos no sería óptimo.

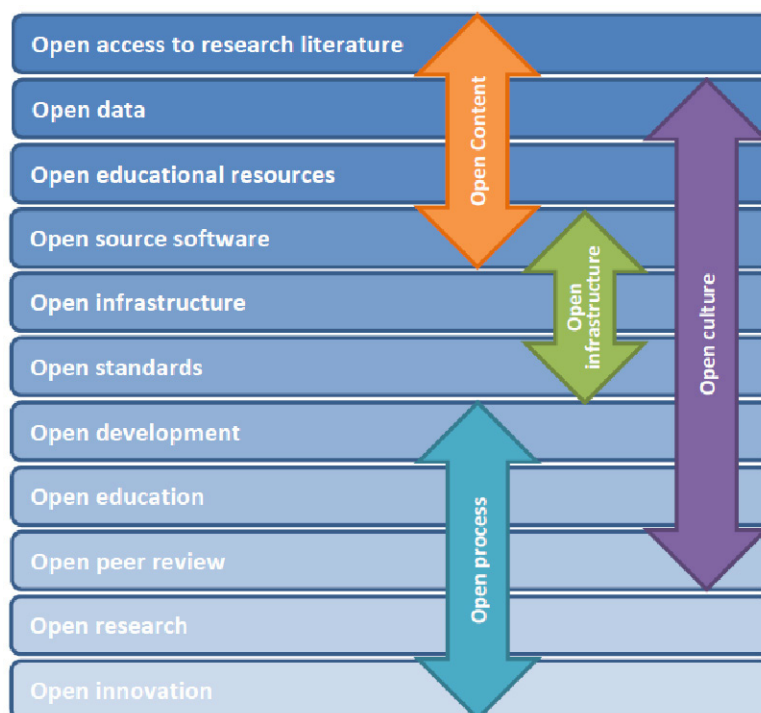
Si se opta por lo abierto como modus operandi, implica que las partes constituyentes del sistema deben trabajar conjuntamente. Tomar esta decisión comporta:

- Un cambio en el sistema de evaluación.
- Abandonar el dogma "*publish or perish*".
- Introducir un sistema de criterios múltiples de evaluación de la producción científica basado en la contribución al avance del conocimiento en lugar del avance de la propia carrera de los investigadores.

Para alcanzar este "estado de abierto" los autores hacen las siguientes recomendaciones:

1. Que las instituciones financiadoras, académicas y de investigación adopten políticas basadas en el abierto como modus operandi para cualquier actividad financiada con fondos públicos.
2. Estas políticas deben incluir procedimientos para el seguimiento de su cumplimiento.
3. La colaboración e implicación de los investigadores debe incentivarse por:
  - Adopción de nuevos sistemas apropiados de evaluación y reconocimiento.
  - Servicios de apoyo con respecto a los derechos de autor y licencias.

**Figura 5.** Rango de “abiertos” (*opens*) de acuerdo a la concepción del acceso abierto como modus operandi (fuente original: e-InfraNet: ‘Open’ as the default modus operandi for research and higher education)



4. Capacitación dirigida al personal de la institución.
5. Asegurarse que la interoperabilidad de los sistemas y servicios sea un componente principal de la e-infraestructura abierta.

Es posible que todavía estemos lejos de alcanzar esta forma de concebir el ciclo de la vida académica, pero no imposible.

#### 4. BIBLIOGRAFÍA

- Bartling, S.; Friesike, S. (2014). Towards Another Scientific Revolution. En Bartling, S.; Friesike, S. (Eds.). *Opening Science*. Cham; Springer International Publishing. doi: <http://dx.doi.org/10.1007/978-3-319-00026-8>
- Borgman, C. (2012a). Why are the attribution and citation of scientific data important? In Uhler, P. F. (Ed.). *For attribution: Developing scientific data attribution and citation practices and standards: Summary of an international workshop*. Washington, D.C.; National Academies Press. pp. 1-10. [http://www.nap.edu/catalog.php?record\\_id=13564](http://www.nap.edu/catalog.php?record_id=13564) [Consultado 28/01/2014].
- Borgman, C.L. (2012b). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, vol. 63(6), pp. 1059-1078. doi: <http://dx.doi.org/10.1002/asi.22634>
- Budapest Open Access Initiative (2012). Diez años desde la *Budapest Open Access Initiative*: hacia lo abierto por defecto. <http://www.budapestopenaccessinitiative.org/boai-10-translations/spanish> [Consultado 28/01/2014].
- COAR (2013). Horizon 2020 – Outline of a Pilot for Open Research Data. [http://www.coar-repositories.org/files/Horizon\\_2020\\_Open\\_Data\\_Pilot\\_20130703\\_final.pdf](http://www.coar-repositories.org/files/Horizon_2020_Open_Data_Pilot_20130703_final.pdf) [Consultado 28/01/2014].
- Coates, H.; Konkiel, S.; Witt, M. (2013). Data Services: Making It Happen. <https://scholarworks.iupui.edu/handle/1805/3278> [Consultado 30/01/2014].
- Costas, R.; Meijer, I.; Zahedi, Z.; Wouters, P. (2013). The Value of Research Data - Metrics for datasets from a cultural and technical point of view. A Knowledge Exchange Report, <http://www.knowledge-exchange.info/datametrics> [Consultado 28/01/2014].
- DataCite (2014). DataCite, re3data.org, and Databib Announce Collaboration <http://www.datacite.org/node/115> [Consultado 10/04/2014].
- Denton Declaration: An Open Access Data Manifesto (2012). *Annual University of North Texas Symposium on Open Access*. <http://openaccess.unt.edu/denton-declaration> [Consultado 28/01/2014].
- Digital.CSIC. Buenas prácticas y directrices para datos de investigación en Digital.CSIC.

- <http://digital.csic.es/politicas/politicaDatos.jsp> [Consultado 30/01/2014].
- e-InfraNet (2013). 'Open' as the Default Modus Operandi for Research and Higher Education. <http://e-infranet.eu/output/e-infranet-open-as-the-default-modus-operandi-for-research-and-higher-education/> [Consultado 28/01/2014].
- European Commission (2009). Open Access in FP7. <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300> [Consultado 30/01/2014].
- European Commission (2013). Commission launches pilot to open up publicly funded research data. [http://europa.eu/rapid/press-release\\_IP-13-1257\\_en.htm](http://europa.eu/rapid/press-release_IP-13-1257_en.htm) [Consultado 30/01/2014].
- FECYT- Grupo de Trabajo de "Depósito y Gestión de datos en Acceso Abierto" del proyecto RECOLECTA (2012). La conservación y reutilización de los datos científicos en España. Informe del grupo de trabajo de buenas prácticas. Madrid: Fundación Española para la Ciencia y la Tecnología, FECYT. <http://digital.csic.es/handle/10261/65317> [Consultado 10/10/2014]
- Harnad S.; Brody T. (2004). Comparing the impact of open access vs. non open access articles in the same journals. *D-Lib Magazine*, vol. 10 (6).
- Hitchcock, S. (2013). The effect of open access and downloads ('hits') on citation impact: a bibliography of studies. <http://eprints.soton.ac.uk/354006/1/oacitation-biblio-snapshot0613.html> [Consultado 28/01/2014].
- MELIBEA. Directorio y estimador de políticas en favor del acceso abierto a la producción científica. <http://www.accesoabierto.net/politicas/> [Consultado 30/01/2014].
- Merlo Vega, J. A.; Angosto Castro, A.; Ferreras Fernández, T.; Gallo León, J. P.; Maestro, J. Á.; Ribes Llopes, I. (2011). Ciencia 2.0: aplicación de la web social a la investigación. Madrid; REBIUN.
- MIT Libraries (2014). Manage Your Data. Data Management: Subject Guides. <http://libraries.mit.edu/guides/subjects/data-management/> [Consultado 28/01/2014].
- National Science Foundation (2013). NSF Data Management Plan Requirements. <http://www.nsf.gov/bfa/dias/policy/dmp.jsp> [Consultado 28/01/2014].
- National Institutes of Health (NIH) (2008). NIH Public Access Policy Details. <http://publicaccess.nih.gov/policy.htm> [Consultado 30/01/2014].
- Odisea. International registry on Research data. <http://odisea.ciepi.org/es> [Consultado 30/01/2014].
- Open Knowledge Foundation (2014a). Definición de conocimiento abierto. <http://opendefinition.org/od/espanol/> [Consultado 11/04/2014].
- Open Knowledge Foundation (2014b). Guide to Open Data Licensing. <http://opendefinition.org/guide/data/> [Consultado 11/04/2014].
- OpenAIRE (Open Access Infrastructure for Research in Europe). <http://www.openaire.eu/> [Consultado 28/01/2014].
- Pampel, H.; Vierkant, P.; Scholze, F.; Bertelmann, R.; Kindling, M.; Klump, J.; Goebelbecker, H.; Gundlach, J.; Schirmbacher, P.; Dierolf, U. (2013). Making Research Data Repositories Visible: The re3data.org Registry. *PLoS ONE*, 8(11): e78080. doi: <http://dx.doi.org/10.1371/journal.pone.0078080>
- Piwowar, H.; Vision, T.J. (2013). Data reuse and the open data citation advantage. PeerJ PrePrints 1:e1v1. doi: <http://dx.doi.org/10.7287/peerj.preprints.1v1>
- Ribeiro, C.; Hernández-San-Miguel, J. (2013). DATABIB: an online catalog of research data repositories. *I Congresso ISKO Espanha e Portugal / XI Congresso ISKO Espanha*. Porto, Portugal.
- Scientific Data. <http://www.nature.com/scientificdata/> [Consultado 28/01/2014].
- SHERPA/JULIET. Juliet Statistics. <http://www.sherpa.ac.uk/juliet/stats.php?la=en&mode=simple> [Consultado 29/01/2014].
- SPARC. HowOpenIsIt? New Guide Released. <http://sparc.arl.org/resource/howopenisit> [Consultado 28/01/2014].
- Swan, A. (2010). The open access citation advantage: studies and results to date. <http://eprints.ecs.soton.ac.uk/18516> [Consultado 26/01/2014].
- The Royal Society (2012). Science as an Open Enterprise. The Royal Society Science Policy Centre report 02/12. <http://royalsociety.org/policy/projects/science-public-enterprise/report> [Consultado 28/01/2014].
- Wagner, B. (2010). Open access citation advantage: An annotated bibliography. *Issues in Science and Technology Librarianship*, vol. 60 (2). doi: <http://dx.doi.org/10.5062/F4Q81B0W>
- Wellcome Trust (2011). Sharing research data to improve public health: full joint statement by funders of health research. <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/WTDV030690.htm> [Consultado 29/01/2014].
- White House Directive (2013). Increasing Access to the Results of Federally Funded Scientific Research. [http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf) [Consultado 30/01/2014].
- Wicherts, J.M.; Bakker, M.; Molenaar, D. (2011). Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results. *PLoS ONE* 6(11): e26828. doi: <http://dx.doi.org/10.1371/journal.pone.0026828>